COMP6714 Course Introduction

Table of Contents

1. Pre-requisites
2. Course philosophy and teaching strategies
3. Course aims
4. Learning outcomes
5. Administrative Components
6. Assessment
7. Reading e-mail
8. Academic honesty and plagiarism
9. CSE's Studuent Conduct Statement
10. Course Material
11. Further information
12. Continual Course Improvement

Units of credit	6
Parallel teaching	no

1. Pre-requisites

The formal pre-requisite for this course is

- For UG, MATH1081 and COMP2041
- For PG, COMP9020 and COMP9021

The knowledge that we assume from these courses is:

- experience with procedural programming, and an understanding of a range of data structures (e.g., trees, graphs, hash-tables) and algorithms (e.g., sorting, divide-and-conquer)
- knowledge of discrete mathematics, including sets, logic, functions and relations, and graphs and trees.

2. Course philosophy and teaching strategies

The learning foci in this course are primarily lectures (theoretical knowledge) and projects (practical knowledge). The course will have an emphasis on problem solving for real applications.

Students will learn the main contents of the course through lectures.

3. Course aims

This course aims to introduce theories and methods of information retrieval, and practical issues

regarding search engines.

The course is designed to be practical. As such, real-life examples of information retrieval will also be used throughout the course.

4. Learning outcomes

Students successfully completing this course will be able to:

- understand the whole process of information retrieval and search engines
- understand various document and retrieval models used in information retrieval
- understand various indexing and query processing tehniques and their variants
- develop solutions for real problems using existing technologies
- appreciate the past, present and future of information retrieval and search engine technologies

The learning outcomes are closely related to UNSW graduate attributes 1 — 6. For example, bonus questions will be given in the assignment in order to encourage independent learning and critical thinking.

5. Administrative Components

See the course hompage (http://www.cse.unsw.edu.au/~cs6714) for (up-to-date) information regarding **Course Staff**, **Course Schedule**, and **Course Resource List**.

As the course is continuously developing together with the information retrieval field, the course schedule is subject to change too. Please read the introduction slides (lecture notes of the first week) for the course schedule in the current offering.

6. Assessment

The assessment of the course comprises of written assignments, programming projects and exams.

```
ass1       : 25%
proj1      : 25%
final exam : 50%
```

Note that you also need to score at least 40 (out of 100) in the final exam to pass the course. Otherwise, your score will be capped at 39FL.

Both written and programming assignments are helpful to achieve deep understanding of the course materials and develop problem solving abilities.

Relationship to the Learning Outcomes:

- All assignments and the final examination are used to assess whether you have achieved the objectives of the subject.
- The programming assignment, in particular, gives you the opportunities to engage in real problems, and helps you to develop in-depth understanding of the course materials and develop problem-solving skills.

Grading Criteria: Grading criteria for each assessment will be detailed in the specification.

Late submission: Assignments/projects submitted late are subject to late penalties, which are specified in the assignment/project specifications; no late submission is allowed if the specification does not specify the late penalty scheme.

Typically "Hard" Late penalties are used in this course, which reduces the actual mark obtained. Thus if the assignment is marked out of 10, and students *A* and *B* hand in assignments worth 9 and 7, both receiving 20% penalty, *A* gets 9 * (1 - 0.2) = 7.2 and *B* gets 7 * (1 - 0.2) = 5.6.

Assignment submission: Assignment submission procedure is described in the assignment specification document, which will be linked to this page when the assignment specification becomes available. Generally assignments are submitted electronically using the give program running on the School's computer systems (in labs, and on servers). Details are in the assignment specifications.

7. Reading e-mail

You should check your school e-mail frequently in case of announcements relating to this course. We assume that you read e-mail sent to your CSE account by the next working day during teaching sessions.

- How to forward your CSE E-mail to another e-mail address
- CSE E-mail guidelines

8. Academic honesty and plagiarism

Copying assignments is unacceptable. Assignments will be checked. The penalties for copying range from receiving no marks for the assignment, through receiving a mark of 00 FL for the course, to expulsion from UNSW (for repeat offenders). Allowing someone to copy your work counts as plagiarism, even if you can prove that it is your work.

Further details of the School plagiarism policy can be found here. (You acknowledged receipt of these rules when you obtained your CSE computer account, and the link above is for your convenience so that you can review the rules now.)

We are aware that a lot of learning takes place in student conversations, and don't wish to discourage those. However, it is important, for both those helping others and those being helped, not to provide/accept any programming language code in writing, as this is apt to be used exactly as is, and lead to plagiarism penalties for both the supplier and the copier of the codes. Write something on a piece of paper, by all means, but tear it up/take it away when the discussion is over.

If you are new to studying in Australia, be aware that attitudes to plagiarism at UNSW may be different from those in your home country. Make sure you are clear about the rules here at UNSW. In brief, and for the purposes of COMP6714, plagiarism includes copying or obtaining all, or a substantial part, of the material for your assignment, whether programming language code, or written or graphical report material, without written acknowledgement in your assignment from:

- location on the internet;
- a book, article or other written document (whether published or unpublished) whether electronic or on paper or other medium;
- another student, whether in your class or another class;
- a non-student (e.g. from someone who writes assignments for money)

Note that if you copy code or other material from another student or non-student with acknowledgement,

you will not be penalised for plagiarism, but you are unlikely to get any marks for the copied material. If you use code found in a publication (on the internet or otherwise) then the marks you get for this will be at the marker's discretion, and will reflect the marker's perception of the amount of work you put into finding and/or adapting the code, and the degree to which you understand the code.

Note also that there is a big difference between being able to understand someone else's code, and writing that code yourself from scratch. A computer programmer has to be able to write code from scratch. The assignments provide opportunities for you to develop the skills necessary to write your own code. Use these opportunities!

9. CSE's Studuent Conduct Statement

The **Student Code of Conduct** (Information, Policy) sets out what the University expects from students as members of the UNSW community. As well as the learning, teaching and research environment, the University aims to provide an environment that enables students to achieve their full potential and to provide an experience consistent with the University's values and guiding principles. A condition of enrolment is that students inform themselves of the University's rules and policies affecting them, and conduct themselves accordingly.

In particular, students have the responsibility to observe standards of equity and respect in dealing with every member of the University community. This applies to all activities on UNSW premises and all external activities related to study and research. This includes behaviour in person as well as behaviour on social media, for example Facebook groups set up for the purpose of discussing UNSW courses or course work. Behaviour that is considered in breach of the Student Code Policy as discriminatory, sexually inappropriate, bullying, harassing, invading another's privacy or causing any person to fear for their personal safety is serious misconduct and can lead to severe penalties, including suspension or exclusion from UNSW.

If you have any concerns, you may raise them with your lecturer, or approach the School Ethics Officer, Grievance Officer, or one of the student representatives.

Plagiarism is defined as using the words or ideas of others and presenting them as your own. UNSW and CSE treat plagiarism as academic misconduct, which means that it carries penalties as severe as being excluded from further study at UNSW. There are several on-line sources to help you understand what plagiarism is and how it is dealt with at UNSW:

- Plagiarism and Academic Integrity
- UNSW Plagiarism Procedure

Make sure that you read and understand these. Ignorance is not accepted as an excuse for plagiarism. In particular, you are also responsible that your assignment files are not accessible by anyone but you by setting the correct permissions in your CSE directory and code repository, if using. Note also that plagiarism includes paying or asking another person to do a piece of work for you and then submitting it as your own work.

UNSW has an ongoing commitment to fostering a culture of learning informed by academic integrity. All UNSW staff and students have a responsibility to adhere to this principle of academic integrity. Plagiarism undermines academic integrity and is not tolerated at UNSW. Plagiarism at UNSW is defined as using the words or ideas of others and passing them off as your own.

If you haven't done so yet, please take the time to read the full text of

- UNSW's policy regarding academic honesty and plagiarism

The pages below describe the policies and procedures in more detail:

- Student Code Policy
- Student Misconduct Procedure
- Plagiarism Policy Statement
- Plagiarism Procedure

You should also read the following page which describes your rights and responsibilities in the CSE context:

- Essential Advice for CSE Students

10. Course Material

- All lecture notes, tutorial and lab notes will be posted on the course web site. These notes summarises the major contents and help you to understand the materials when you read the textbook later. You definitely need to read the corresponding chapters in the textbook to gain a full understanding of the course materials.

```
[MRS08] Christopher D. Manning, Prabhakar Raghavan and Hinrich
Schutze, Introduction to Information Retrieval, Cambridge
University Press. 2008.


[CMS09] W. Bruce Croft, Donald Metzler, and Trevor Strohman,
Search Engines: Information Retrieval in Practice. Pearson. 2009.


[JM19] Dan Jurafsky, James H. Martin, Speech and Language Processing (3rd ed. draft). 2019.
```

- Reference books for this course are:

```
[BCC10] Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack,
Information Retrieval: Implementing and Evaluating Search Engines,
The MIT Press. 2010.


[BB99] Ricardo Baeza-Yates and Ber thier Ribeiro-Neto, Modern Information
Retrieval, Addison Wesley. 1999.
```

- Useful Resources
  - http://nlp.stanford.edu/IR-book/information-retrieval.html
  - http://www.ir.uwaterloo.ca/book/

11. Further information

- Students enrolled in the course are expected to attend all classes
- The use of School of Computer Science and Engineering computing laboratories is subject to rules described in the Yellow Form, which you acknowledge (electronic) receipt of when you receive your computing account. The Yellow Form also outlines what to do in case illness or misadventure that affects your assessment, and supplementary examinations procedures within the School of Computer Science and Engineering.
- Information on UNSW Occupational Health and Safety policies and expectations

- Equity and Diversity issues

12. Continual Course Improvement

Each year feedback is sought from students and other stakeholders about the courses offered in the School and continual improvements are made based on this feedback. UNSW's Course and Teaching Evaluation and Improvement (CATEI) Process (http://www.unsw.edu.au/learning/pve/catei.html) is one of the ways in which student evaluative feedback is gathered. Significant changes to courses and programs within the School are communicated to subsequent cohorts of students.

There were no serious issues noted in the previous evaluation. Thus, we will only make incremental refinements to preserve the already high standard of the course.

Wei Wang

Last updated 2020-09-06 02:04:05 AEDT